Compound V2 Wallet Scoring: Methodology

Author: Sai Satyam Jena
Date: 5th May 2025

1. Executive Summary

This document details the methodology developed to assign a credit score (0-100) to Compound V2 protocol wallets, based solely on their historical transaction behavior. Faced with raw, unlabeled transaction data, the core challenge was to define and quantify "creditworthy" behavior within the DeFi lending context. An unsupervised approach was employed, focusing on feature engineering derived directly from transaction logs (deposits, withdraws, borrows, repays). Key behavioral indicators like repayment history, collateralization proxies, leverage management, and user lifespan were quantified. A transparent, weighted feature sum model was implemented to aggregate these indicators into a final score, where higher scores reflect responsible, stable participation, and lower scores indicate potentially risky or extractive behavior. The resulting system provides an interpretable, data-driven assessment of wallet quality based on observed on-chain actions within the provided dataset.

---

2. Data Source & Preprocessing

- **Data Source:** Raw transaction-level data from the Compound V2 protocol on Ethereum, provided as a set of JSON files. The three **largest files (compoundV2_transactions_ethereum_chunk_0.json,**

  compoundV2_transactions_ethereum_chunk_1.json,

  compoundV2_transactions_ethereum_chunk_2.json)

were selected for analysis to ensure a significant representation of protocol activity.

- **Initial Loading & Structure:** Analysis revealed the data was structured as nested JSON, with top-level keys (e.g., deposits, borrows) mapping to lists of individual transaction records. Initial attempts using standard Pandas JSON loading failed on some files due to inconsistent list lengths between keys. A robust custom loading process was implemented using Python's **json** library to parse each file, iterate through known transaction type keys, and extract individual transaction records into a unified list. This ensured all valid transactions from the selected files (approx. 120,000) were captured.
- **Flattening:** The primary data transformation involved flattening the nested structure into a tabular format where each row represents a single transaction event. Key fields extracted included:
    - **wallet_address**: The interacting wallet ('account.id' from source).
    - **timestamp:** UNIX timestamp of the transaction.
    - **std_tx_type:** Standardized transaction type ('Deposit', 'Withdraw', 'Borrow', 'Repay') mapped from source keys.
    - **amount:** Raw token amount involved.
    - **amountUSD:** USD value of the transaction amount at the time (crucial for cross-asset comparison).
    - **asset_symbol**: The token involved (e.g., 'ETH', 'DAI').
    - **tx_hash**: Unique identifier for the transaction record.
- **Data Cleaning & Type Conversion:**
    - ***Timestamps:*** Converted from UNIX seconds to standard datetime objects.
    - ***Numeric Values:*** amount and amountUSD were converted to float64, ensuring accurate numerical representation. Fortunately, no missing amountUSD values were encountered after conversion.

  - ***Wallet Addresses:*** Standardized to lowercase for consistency.
- **Key Observation:** During initial exploration, it was noted that **no 'Liquidation' events** were present within the processed data segments. This significantly influenced the feature engineering strategy, requiring a focus on proxy measures for risk rather than direct liquidation counts.

---

3. Defining Wallet Behavior

In the absence of predefined labels, defining "good" **(creditworthy)** and "bad" **(risky/unreliable)** behavior is the foundational step. Within the context of a DeFi lending protocol like Compound, behavior can be assessed based on its contribution to protocol health, stability, and adherence to the implicit contract of borrowing and lending.

- **Principles of "Good" / Creditworthy Behavior:**
  - ***Reliable Repayment****:* Consistently meeting debt obligations is the cornerstone of creditworthiness. Users who borrow and fully repay demonstrate responsibility.
  - ***Stable Participation & Collateralization****:* Long-term engagement and maintaining sufficient collateral (net positive deposits over time) contribute to protocol liquidity and stability. These users are less likely to be transient or purely extractive.
  - ***Prudent Leverage Management****:* Borrowing well within collateral limits indicates sound risk management and reduces systemic risk.
  - ***Value Contribution****:* Providing significant net value (in USD terms) to the protocol through deposits suggests a commitment beyond immediate borrowing needs.
- **Principles of "Bad" / Risky Behavior:**
  - ***Failure to Repay****:* Low or zero repayment of borrowed amounts is a primary indicator of poor creditworthiness or default risk.

- ***Excessive Leverage:*** Borrowing heavily relative to deposited collateral (high borrow-to-deposit ratio) signifies high risk-taking, increasing the potential for losses (even if liquidations weren't observed in *this* data).
- ***Ephemeral Interaction:*** Very short participation lifespans, especially combined with high leverage or minimal repayment, can suggest hit-and-run strategies, potential exploits, or bot activity not aligned with sustainable protocol use.
- ***Negative Net Value:*** Consistently withdrawing more value than deposited, especially while borrowing, can be extractive.

- **Adapting to Data Constraints:** Given the absence of liquidation data, the focus shifted heavily towards **quantifying *repayment behavior*** and ***leverage proxies*** as the **primary indicators of risk and responsibility**.

---

4. Feature Engineering: Quantifying Behavior

The defined behavioral principles were translated into quantitative features calculated for each unique wallet_address, using amountUSD for value normalization:

- **Lifespan & Activity:**
  - **wallet_lifespan_days:**

    (last_seen_timestamp - first_seen_timestamp).days + 1.

    Intuition: Measures the duration of engagement. Longer lifespans generally suggest more stable, committed users.
  - **total_transactions:** Total count of interactions.

    Intuition: Basic measure of activity volume.

- **transactions_per_day:**

  total_transactions / wallet_lifespan_days.

  Intuition: Measures activity frequency. Very high values might indicate bot activity; very low values indicate inactivity.

- **Value & Collateralization:**
  - **Deposit_amountUSD_sum**, **Withdraw_amountUSD_sum**: Total USD value deposited and withdrawn.
  - **net_deposit_amountUSD**:

    Deposit_amountUSD_sum - Withdraw_amountUSD_sum.

    Intuition: Represents the net USD value the wallet has contributed or withdrawn. A positive value indicates net collateral provision. Strong indicator of user stake.

- **Borrowing & Repayment Behavior:**
  - **Borrow_amountUSD_sum**, **Repay_amountUSD_sum**: Total USD value borrowed and repaid.
  - **has_borrowed**: Binary flag (1 if Borrow_count > 0, else 0).
  - **repayment_ratio_usd**: Repay_amountUSD_sum / (Borrow_amountUSD_sum + epsilon). Capped [0, 1].

    Intuition: **Core creditworthiness indicator.** Measures the proportion of borrowed USD value repaid. Values near 1.0 signify responsible repayment. Values near 0.0 indicate failure to repay borrowed value.

  - **borrow_to_deposit_ratio_usd**: Borrow_amountUSD_sum / (net_deposit_amountUSD + epsilon). Capped [0, 10], with special handling for borrowing with <=0 net deposit (assigned 10.0)

    Intuition*:* **Core risk indicator proxying leverage.** This ratio quantifies the total USD value borrowed relative to the net USD value the wallet has committed via deposits minus withdrawals. When net deposits are positive, it reflects standard leverage. However, borrowing while having zero or negative net deposits signifies taking on debt without a net collateral stake, a fundamentally higher risk. The model assigns a fixed maximum value (10.0) to represent this high-risk state consistently with extremely high leverage against positive collateral. Higher values strongly signal higher risk.

- **Engagement:**
  - **unique_assets_interacted**: Count of distinct asset symbols used.

    Intuition: Suggests broader engagement with the protocol's offerings.

---

5. Scoring Model Design & Rationale

- **Model Choice:** A **Weighted Feature Sum** model was selected.

  Rationale:

  - ***Interpretability:*** Essential for understanding *why* a wallet receives a certain score, directly linking features to the final output.
  - ***Unsupervised Context:*** Ideal when ground truth labels are unavailable. Allows encoding domain knowledge (our behavior definitions) directly into the model via weights.
  - ***Transparency:*** The scoring logic is explicit and reproducible.
  - ***Appropriateness:*** Avoids unnecessary complexity; sophisticated models are hard to train and justify without labels and may not provide more meaningful results for this task.

- **Normalization:** All selected features were scaled to a [0, 1] range using **MinMaxScaler**.

  Rationale: This ensures features with different natural scales (e.g., lifespan days vs. ratios) contribute proportionally based on their assigned weights, not just their magnitude.

- **Weight Assignment:** Weights were assigned to the normalized features based on their perceived importance in reflecting the defined principles of good/bad behavior, informed by observed feature distributions:

- **repayment_ratio_usd**: **+0.30** (Highest positive weight: Strongest indicator of responsible behavior)
- **net_deposit_amountUSD**: **+0.25** (High positive weight: Rewards collateral provision and value contribution)
- **wallet_lifespan_days**: **+0.20** (Moderate positive weight: Rewards stability and commitment)
- **unique_assets_interacted**: **+0.05** (Low positive weight: Slightly rewards broader engagement)
- **borrow_to_deposit_ratio_usd**: **-0.15** (Significant negative weight: Penalizes high leverage/risk proxy)
- **transactions_per_day**: **-0.05** (Low negative weight: Slightly penalizes potentially anomalous high frequency)

- **Score Calculation:** The final score was calculated as:
  Raw Score = Σ (normalized_feature * weight)
  This Raw Score was then scaled again using MinMaxScaler to produce the final score neatly within the 0-100 range for interpretability.

---

6. Outputs Generated

The designed Python script processes the data and calculates features and scores as described. Upon completion, the script generates **two key output files**:

1. **top_1000_scores.csv**

   Contains the **wallet_address** and **final_score** for the **1,000 highest-scoring wallets**, **sorted in descending order by score**. **This serves as the primary deliverable for identifying top-performing wallets.**

2. **all_wallet_scores_and_features.csv**

   Contains all **calculated features** and scores for ***every* unique wallet** processed from the **input data**, also sorted by **final_score**. **This comprehensive file is used for detailed analysis, including the**

examination of both high-scoring and low-scoring wallets presented in the accompanying Wallet Analysis document.

---

7. Limitations & Future Work

- **Limitations:**
  - ***Data Coverage****:* Analysis limited to the three provided data chunks; behavior might differ in other periods. Absence of liquidation data required reliance on proxy measures.
  - ***Feature Simplicity****:* Features are aggregate summaries; they don't capture intra-day volatility or time-series patterns (e.g., consistency of deposits).
  - ***Weight Subjectivity****:* While based on first principles, the assigned weights inherently involve some expert judgment.
  - ***Bot Detection****:* transactions_per_day is a very basic heuristic for bot detection.
  - ***External Factors****:* Scoring ignores external wallet activity or market conditions not present in the transaction logs.
- **Future Work:**
  - Incorporate liquidation data if available for a more direct risk measure.
  - Develop time-series features (e.g., rolling volatility of net deposits, time between borrows/repays).
  - Explore more sophisticated anomaly/bot detection techniques.
  - Experiment with unsupervised clustering (e.g., K-Means on features) to identify distinct behavioral archetypes as a complementary analysis.
  - Sensitivity analysis on feature weights.

---

8. Conclusion

This methodology successfully produces an interpretable, behavior-based credit score (0-100) for Compound V2 wallets using only raw transaction data**. By defining creditworthiness from first principles within the DeFi context** and engineering relevant features (particularly repayment and leverage proxies), **the weighted scoring model effectively differentiates wallets based on their observed historical activity**. While acknowledging data limitations and weight subjectivity, the resulting scores provide a valuable, justifiable starting point for assessing wallet quality and risk.